

Practical 1: ANN Training and Prediction

Dr. Wenjun Wang

8th Sep 2023

Due to 17:00, 22nd Sep 2023

In this practical project, you will be asked to tackle two tasks and then submit your work in *no more than 3 pages*, with *codes* and *output* document (optional) attached.

Task1: Manual Training an ANN Classifier

To construct a simple perceptron to and train it by stochastic gradient method on the following dataset.

Tip:

(1) the perceptron should have 3 variable input and 1 constant (bias) input.

(2) use sigmoid function as activation function, $\sigma(x) = \frac{1}{1+e^{-x}}$, and $\sigma(x)' = \sigma(x) \times [1 - \sigma(x)]$.

| ID | Study Hrs per Week | Sleep Hrs per Week | Quiz Marks | Exam Marks |
|----|--------------------|--------------------|------------|------------|
| 1 | 24 | 56 | 78 | pass |
| 2 | 5 | 66 | 52 | fail |
| 3 | 92 | 42 | 82 | pass |
| 4 | 31 | 63 | 67 | pass |
| 5 | 15 | 70 | 44 | fail |
| 6 | 12 | 45 | 35 | fail |

Please report the details of calculation of at least 3 epochs, if it is converged, why? If it not, why, and what is your suggestion to get an optimal solution.

Please also report your final MSE and the predictor model.

Task2: Churn Rate Prediction and Analysis

To predict churn rate of customer with one of a European bank by using ANN.

Please see the demo in Blackboard, then use the trained model for prediction.

Please ask the following questions, according to your calculations:

(1) Will the customer leave the bank or stay with it, please give the probability of leave? Given information are list as follows:

| Geography | Credit Score | Gender | Age | Tenure | Balance | No. of Products | credit card | Active | Estimated Salary |
|-----------|--------------|--------|-----|---------|----------|-----------------|-------------|--------|------------------|
| France | 600 | Male | 40 | 3 years | \$ 60000 | 2 | yes | yes | \$ 50000 |

(2) Given a threshold of 0.5, please print out the binary result.

(3) Try to get a confusion matrix and accuracy on test data set.

(4) Do you think the normalization of data is important, or not? Please verify your result and give evidence for supporting.

Supplementary:

Please see dataset attached named as churn_modelling.csv. The description of items recorded are as following:

| Customer Id | Surname | CreditScore | Geography | Gender | Age | Tenure | Balance | NumOfProducts | HasCrCard | IsActiveMember | EstimatedSalary | Exited |
|-------------|---------|-------------------|-----------|----------------|----------|---------------------------|-----------------------|--|----------------------------|---|-----------------|---|
| ID | Name | Evaluated by Bank | Country | Male or female | By years | Years that with this bank | Current money balance | e.g. has a debit card and a credit card, then this number =2 | 1 means yes and 0 means no | Dose the usere visit bank account or pay by this card in recent 6 month? Yes 1; no 0. | Done by Bank | Does the user leave or not in last 6 month. Yes 1; No,0 |